



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/978,192	10/15/2001	Avi J. Ashkenazi	GNE.2630P1C9	3437

7590 05/20/2004  
Ginger R. Dreger, Esq.  
Knobbe Martens Olson & Bear  
620 NEWPORT CENTER DRIVE  
SIXTEENTH FLOOR  
NEWPORT BEACH, CA 92660

EXAMINER

O HARA, EILEEN B

ART UNIT PAPER NUMBER

1646

DATE MAILED: 05/20/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	<b>Application No.</b> 09/978,192	<b>Applicant(s)</b> ASHKENAZI ET AL.	
	<b>Examiner</b> Eileen O'Hara	<b>Art Unit</b> 1646	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☐ Responsive to communication(s) filed on \_\_\_\_.
- 2a) ☐ This action is **FINAL**.      2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 58-63 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 58-63 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 15 October 2001 is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)  | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. ____. |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)   | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)             |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date <u>02/02 &amp; 05/02</u> . | 6) <input type="checkbox"/> Other: ____.  |

### **DETAILED ACTION**

1. Claims 58-63 are pending in the instant application. Claims 1-57 have been canceled and claims 58-63 have been added as requested by Applicant in the Preliminary Amendment filed October 15, 2001.

#### ***Specification***

2. The disclosure is objected to because it contains embedded hyperlinks and/or other form of browser-executable code. See page 124, line 37, page 127, line 18, page 233, line 1, page 175, line 1, page 276, line 1, page 309, line 32, page 311, line 33, page 313, lines 5, 6, 20 and 23, at least. Applicant is required to delete the embedded hyperlink and/or other form of browser-executable code. See MPEP § 608.01.

#### ***Double Patenting***

3. Applicant and the assignee of this application are required under 37 CFR 1.105 to provide the following information that the examiner has determined is reasonably necessary to the examination of this application. 37 CFR 1.78(b) provides that when two or more applications filed by the same applicant contain conflicting claims, elimination of such claims from all but one application may be required in the absence of good and sufficient reason for their retention during pendency in more than one application.

A sequence search of the pending and published application databases has revealed that there are a series of applications in which SEQ ID NO: 7 is present but that do not claim the polypeptide. However, there is at least one other application filed by the applicants which contains the polypeptide of SEQ ID NO: 16 which is identical to the polypeptide of SEQ ID NO: 7, and which may contain possible conflicting claims. Due to the large number of applications that contain this sequence, the examiner is unable to determine if any of these applications have claims directed to this polypeptide. Applicant is required to point out to the Examiner all double patenting issues. See MPEP § 1.105.

The applicant is reminded that the reply to this requirement must be made with candor and good faith under 37 CFR 1.56. Where the applicant does not have or cannot readily obtain an item of required information, a statement that the item is unknown or cannot be readily obtained will be accepted as a complete reply to the requirement for that item.

This requirement is an attachment of the enclosed Office action. A complete reply to the enclosed Office action must include a complete reply to this requirement. The time period for reply to this requirement coincides with the time period for reply to the enclosed Office action.

#### ***Formal Matters***

4. The deposit of biological organisms is considered by the Examiner to be necessary for enablement of the current invention (see MPEP Chapter 2400 and 37 C.F.R."1.801-1.809). Examiner acknowledges the deposit of organisms under accession number ATCC 209786 under terms of the Budapest Treaty on International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure in compliance with this requirement (see specification, pages 372-374).

***Advisory Information***

5. The claims are interpreted such that the fragment of the antibody must also bind the protein. If Applicants intend otherwise, it is suggested the claims be amended to clarify this.

***Claim Rejections - 35 USC § 101 and § 112***

35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

6.1 Claims 58 and 63 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. Claim 58 is directed to an antibody that binds to the polypeptide of SEQ ID NO: 7, and such an antibody could exist in nature. The rejection would be withdrawn if the word “isolated” were inserted in front of “antibody”.

6.2 Claims 58-63 are rejected under 35 U.S.C. 101 because the claimed invention is not supported by either a specific and substantial asserted utility or a well established utility. Claims 58-63 are drawn to antibodies to the protein of SEQ ID NO: 7, identified as PRO274. The instant specification discloses that PRO274 is a 492 amino acid protein, and is presumably a membrane-bound protein with extracellular domain from amino acids 1-85, and transmembrane domains from amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374, 408-423 and 431-445 (Fig. 4). The specification teaches that PRO274 has homology to the 7-transmembrane receptor family and to FN54 protein. However, the protein (or encoding nucleic acids) do not have any specific and substantial utility, or a well established utility, as determined according to

the current Utility Examination Guidelines, Federal Register, Vol. 66, No. 4, pages 1092-1099, Friday, January 5, 2001.

The claims are directed to antibodies to the polypeptide of SEQ ID NO: 7. The specification contains numerous asserted utilities for the polypeptide and encoding nucleic acid at pages 190-199, including use as hybridization probes, in chromosome and gene mapping, in the generation of anti-sense RNA and DNA, to identify molecules that bind to PRO (including agonists and antagonists), to make “knock-out” mice or other animals, in gene therapy, as molecular weight markers, therapeutic agents, and for the production of antibodies. Asserted utilities for the antibodies include diagnostic assays for PRO, *e.g.* detecting its expression in specific cells, tissues or serum, and affinity purification of PRO274. The utilities that pertain solely to nucleic acids (*e.g.* hybridization, chromosome and gene mapping, anti-sense) would not convey to the encoded protein. With respect to the remaining utilities, none of these asserted utilities is specific for the disclosed PRO274 protein or antibodies, as each of the aforementioned utilities could be asserted for any naturally occurring protein and associated antibodies, and further, as none of the asserted utilities requires any feature or activity that is specific to the disclosed PRO274.

The specification teaches that PRO274 has (unspecified) homology to the 7-transmembrane receptor family. The amino acid sequence of the putative PRO274 peptide is shown in Figure 4 of the specification, in which transmembrane domains are identified, however there is no disclosure that the protein is expected to be a transmembrane protein other than identification of putative transmembrane domains. There is no biological activity, expression pattern, phenotype, disease or condition, ligand, binding partner, or any other specific feature

Art Unit: 1646

that is disclosed as being associated with PRO274. Without any information as to the specific properties of PRO274, the mere identification of such as having significant sequence homology to the 7-transmembrane receptor family and/or FN54 protein (which has no disclosed activity) sufficient to impart any particular utility to the polypeptides or the claimed antibodies.

The specification at pages 331-346 describes experiments in which PRO274 encoding genes are asserted to be amplified in the genome of certain human lung primary tumors. At pages 119-137 it is disclosed that nucleic acids encoding PRO274 had  $\Delta C_t$  values of at 1.24, 1.00 and 1.61 for 3 primary lung tumors, LT4, LT16 and LT18, respectively, for which a value of 1.00 corresponds an amplification of two-fold over normal tissue. From Table 8 on page 338, PRO274 was amplified in one (LT4) out of nine human lung tumor adenocarcinoma tumors, and was amplified in two (LT16 and LT18) out of nine human lung tumor squamous cell carcinomas. Given that PRO274 was amplified in only a very small number of tumors of the same type, the data do not support the implicit conclusion of the specification that PRO274 shows a positive correlation with lung cancer, much less that the levels of PRO274 would be diagnostic of such. Cancerous tissue is known to be aneuploid, that is, having an abnormal number of chromosomes (see Sen, 2000, Curr. Opin. Oncol. 12:82-88). The data presented in the specification were not corrected for aneuploidy. A slight amplification of a gene does not necessarily mean overexpression in a cancer tissue, but can merely be an indication that the cancer tissue is aneuploid. Even if the data were corrected for aneuploidy, one of ordinary skill in the art would not conclude that PRO274 would be diagnostic for lung cancer, due to the lack of overexpression in the majority of primary tumor types.

Even *if* the data demonstrated a increase in copy number of PRO274 nucleic acids in primary tumors, such would not be indicative of a use of antibodies to the encoded polypeptide as a diagnostic agent. The preliminary data were not supported by analysis of mRNA or protein expression, for example. Also, it does not necessarily follow that an increase in gene copy number results in increased gene expression and increased protein expression, such that antibodies would be useful diagnostically or as a target for cancer drug development. For example, Pennica et al. (1998, PNAS USA 95:14717-14722) teach that

“An analysis of *WISP-1* gene amplification and expression in human colon tumors showed a correlation between DNA amplification and overexpression, whereas overexpression of *WISP-3* RNA was seen in the absence of DNA amplification. In contrast, *WISP-2* DNA was amplified in the colon tumors, but its mRNA expression was significantly reduced in the majority of tumors compared with the expression in normal colonic mucosa from the same patient.”

See page 14722, second paragraph of left-hand column; pp.14720-14721; Pages 14720-14721, “Amplification and Aberrant Expression of *WISPs* in Human Colon Tumors”.

Gygi et al. (Molecular and Cellular Biology, March 1999, p. 1720-1730), studied over 150 proteins relatively homogeneous in half-life and expression level, and found no strong correlation between protein and transcript levels; for some genes, equivalent mRNA levels translated into protein abundances which varied by more than 50-fold. Gygi et al. concluded that the protein levels cannot be accurately predicted from the level of the corresponding mRNA transcript (abstract and Figure 5).

Thus, the data do not support the implicit assertion that polypeptide of PRO274 polypeptide can be used as a cancer diagnostic. Significant further research would have been required of the skilled artisan to determine whether PRO274 is overexpressed in any cancer to

Art Unit: 1646

the extent that antibodies to the protein could be used as a cancer diagnostic, and thus the implicitly asserted utility is not substantial.

The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

7. Claims 58-63 are also rejected under 35 U.S.C. 112, first paragraph. Specifically, since the claimed invention is not supported by either a specific and substantial asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

8. Claims 58-63 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

8.1 Claim 61 is indefinite because an antibody can't be both an antibody and an antibody fragment.

8.2 Claims 58-63 are indefinite because claim 58 encompasses an antibody that binds to a polypeptide of SEQ ID NO: 7, and claim 63 encompasses an antibody that "specifically binds".

The specification does not define the term "specifically binds" the polypeptide, and it is not clear what this means, and it is not clear what the difference in scope between "binds" and

Art Unit: 1646

“specifically binds” is. Additionally, it is not clear if Applicants intend the art accepted definition of “specifically binds”, that is the antibody binds above background, or alternatively, that the antibody binds exclusively to the protein of SEQ ID NO: 7. If Applicants intend the latter, a rejection under 35 U.S.C. 112, first paragraph for enablement would be made over the claims.

### ***Priority Determination***

35 U.S.C. § 120 states that:

An application for patent for an invention disclosed in the manner provided by the first paragraph of section 112 of this title in an application previously filed in the United States, or as provided by section 363 of this title, which is filed by an inventor or inventors named in the previously filed application shall have the same effect, as to such invention, as though filed on the date of the prior application, if filed before the patenting or abandonment of or termination of proceedings on the first application or on an application similarly entitled to the benefit of the filing date of the first application and if it contains or is amended to contain a specific reference to the earlier filed application.

35 U.S.C. § 119(e) states that:

An application for patent filed under section 111(a) or section 363 of this title for an invention disclosed in the manner provided by the first paragraph of section 112 of this title in a provisional application filed under section 111(b) of this title, by an inventor or inventors named in the provisional application, shall have the same effect, as to such invention, as though filed on the date of the provisional application filed under section 111(b) of this title, if the application for patent filed under section 111(a) or section 363 of this title is filed not later than 12 months after the date on which the provisional application was filed and if it contains or is amended to contain a specific reference to the provisional application.

9. Applicant is advised that the instant application can only receive benefit under 35 U.S.C. § 120 or § 119(e) from an earlier application which meets the requirements of 35 U.S.C. § 112, first paragraph, with respect to the now claimed invention. Because the instant application does not meet the requirements of 35 U.S.C. § 112, first paragraph, for those reasons given above and it is a continuation of the applications listed in the priority map filed June 21, 2002, the prior applications do not meet those requirements and, therefore, are unavailable under 35 U.S.C. §

Art Unit: 1646

120 or § 119(e). The effective priority date of the instant application is considered to be the filing date of this application, October 15, 2001, because the claimed invention is not supported by either a specific and substantial utility or a well established utility.

***Rejections over Prior Art***  
***Claim Rejections - 35 USC § 102***

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

10. Claims 58-63 are rejected under 35 U.S.C. 102(b) as being anticipated by Ho et al., Science, Vol. 289, July 14, 2000, pages 265-270.

Claims 58 and 63 are drawn to antibodies to the polypeptide of SEQ ID NO: 7

Ho et al. disclose mouse and human ANK protein, the human protein of which is 100% identical to the protein of SEQ ID NO: 7 of the instant application (Fig. 3A and attached sequence alignment), and mouse protein which differs from the human protein in only 9 amino acids. Ho et al. made polyclonal antibodies to different epitopes of the the mouse protein, and the antibodies were used in Western blots and in indirect immunofluorescence on transfected cells. Because of the high degree of sequence similarity between the human and mouse proteins, the anti-mouse ANK protein antibodies would also bind to the human protein disclosed by Ho et al. and the protein of SEQ ID NO: 7 of the instant invention. Therefore, Ho et al. anticipates the claims.

***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

9. Claims 59-62 are rejected under 35 U.S.C. 103(a) as being unpatentable over Ho et al., Science, Vol. 289, July 14, 2000, pages 265-270, in view of Immunobiology, The Immune System in Health and Disease, Third Edition, Janeway, And Travers, Ed., 1997.

The teachings of Ho et al. are discussed above. Ho et al. does not teach that the antibodies to ANK protein may be monoclonal, humanized, antibody fragment or labeled.

Immunobiology teaches that antisera have certain disadvantages that relate to the heterogeneity of the antibodies they contain such as differences between batches, production in limited volumes, and thus it is impossible to use the identical serological reagent in a long or complex series of experiments or clinical tests. Antisera also may include antibodies that give unexpected cross-reactions. Immunobiology teaches that monoclonal antibodies can overcome

Art Unit: 1646

these disadvantages with an unlimited supply of antibody molecules of homogeneous structure and known specificity (section 2-10). Immunobiology also teaches that fragments of antibodies called single-chain Fv (Fragment variable) may become valuable therapeutic agents because of their small size, allowing ready tissue penetration (page 3:4-3:5), teach that monoclonal antibodies may become humanized by grafting the antigen-binding loops or CDRs of a mouse monoclonal antibody onto the framework of a human immunoglobulin, a process known as humanization, resulting in antibodies that bind the same antigen as the mouse antibody but are far less immunogenic, and thus can be used for treatment of humans with far less risk of anaphylaxis (section 13-7). Immunobiology also teaches that antibodies (for example to a tumor antigen) can be conjugated to a label such as a radioisotope, which can be used to concentrate the radioactive source to a tumor site and can kill the tumor cells, providing an effective cancer immunotherapy, or antibodies can be conjugated to enzymes or radioisotopes and used in ELISA or RIA binding assays, allowing antigen in unknown samples to be measured easily and rapidly (section 2-7).

It would have been *prima facie* obvious to the person of ordinary skill in the art at the time the invention was made to use make monoclonal, humanized, antibody fragment or labeled antibodies to the ANK protein of Ho et al., for the reasons explained in Immunology, in order to either purify or further study the protein or for therapeutic applications, since the protein is involved in arthritis. The skilled artisan would be motivated to do so since the advantages of such are discussed in Immunology, and there would be a reasonable expectation of success, since making and using these antibodies have been widely and successfully used in the field of immunology.

Art Unit: 1646

***Conclusion***

10. No claim is allowed.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Eileen B. O'Hara, whose telephone number is (571) 272-0878. The examiner can normally be reached on Monday through Friday from 10:00 AM to 6:30 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Yvonne Eyler can be reached at (571) 272-0871.

The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Any inquiry of a general nature or relating to the status of this application should be directed to the Group receptionist whose telephone number is (571) 272-1600.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, se <http://pair-direct.ispto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll free).

Eileen B. O'Hara, Ph.D.



Patent Examiner

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2004, 13:50:23 ; Search time 18 Seconds  
(without alignments)  
1423.251 Million cell updates/sec

Title: US-09-978-188a-7  
Perfect score: 2527

Sequence: 1 MVKFPALHTWPLRLPLVPL.....TDMPEETVDTIVREENE 492

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2527	100.0	492	1 ANKH_HUMAN	Q9HJ11 homo sapien
2	2502	99.0	492	1 ANKH_MOUSE	O9JH22 mus musculu
3	2489	98.5	492	1 ANKH_RAT	P58366 rattus norv
4	2317	91.7	492	1 ANKH_XENLA	P58367 xenopus lae
5	2158.5	85.4	501	1 ANKH_BRARE	P58368 brachydanio
6	1528.5	60.5	355	1 ANKH_TETNG	P58369 tetradodon n
7	114	4.5	416	1 LACY_KLEBOX	P18817 klebsiella
8	113.5	4.5	556	1 NU2M_PODAN	P15578 podospora a
9	111	4.4	622	1 COX1_BACSU	P24010 bacillus su
10	109.5	4.3	610	1 PT2A_ARATH	P46031 arabidopsis
11	107	4.2	422	1 EXU1_BACSU	O34456 bacillus su
12	107	4.2	533	1 LAT2_RAT	O9WV76 rattus norv
13	106	4.2	334	1 YB77_METJA	O58575 methanococc
14	105	4.2	401	1 LSG1_HAEIN	P71399 haemophilus
15	105	4.2	676	1 HPPI_MERAC	O6T193 methanosaar
16	103.5	4.1	398	1 BCR_HAEIN	P45123 haemophilus
17	103	4.1	417	1 LACY_ECOLI	P02920 escherichia
18	102.5	4.1	446	1 CITN_SALTY	P31603 salmonella
19	102.5	4.1	531	1 LAT2_MOUSE	O9QXW9 mus musculu
20	102.5	4.1	533	1 MYIN_RHTR	O05457 rhizobium t
21	102.5	4.1	641	1 NUSM_AITMA	P50365 allomyces m
22	101.5	4.0	446	1 CITN_SALDU	P31603 salmonella
23	100	4.0	982	1 YS96_CAEEL	O09965 caenorhabdi
24	99	3.9	385	1 Y421_METTH	O26521 methanobact
25	99	3.9	535	1 LAT2_HUMAN	O35140 neosporea
26	98.5	3.9	583	1 NU2M_NEUCR	P38820 saccharomyc
27	98.5	3.9	946	1 YB76_YEAST	P39961 saccharomyc
28	98	3.9	480	1 YEH4_YEAST	P37555 bacillus su
29	98	3.9	532	1 YABM_BACSU	P79235 ponzo pygma
30	97.5	3.9	346	1 FMR1_POMPY	O81847 drosophila
31	97.5	3.9	3803	1 TRAI_DROME	P37482 bacillus su
32	97	3.8	402	1 YICB_BACSU	P79176 gorilla gor
33	96.5	3.8	346	1 FMR1_GORGO	

34	95.5	3.8	416	1 LACY_CITFR	P47224 citrobacter
35	95.5	3.8	483	1 NORM_VIBVU	O8d9n8 vibrio vuln
36	95	3.8	457	1 NORM_ECO57	P58164 escherichia
37	95	3.8	459	1 NU4M_BALPH	P24975 balaenopter
38	95	3.8	460	1 NU4M_ASTPE	P11992 asterina pe
39	95	3.8	464	1 NORM_HAEIN	P45272 haemophilus
40	95	3.8	501	1 LYSI_CORGL	P35865 corynebacte
41	95	3.8	511	1 ALG8_DROME	O9W3V8 drosophila
42	95	3.8	654	1 NUSM_RHIST	P50367 rhizopus st
43	94.5	3.7	507	1 TT12_ARATH	O91Y33 arabidopsis
44	94.5	3.7	518	1 SP58_BACSU	O00758 bacillus su
45	94	3.7	421	1 VG2_BP1KE	P03660 bacterioph

ALIGNMENTS

RESULT 1  
ID ANKH\_HUMAN STANDARD; PRT; 492 AA.  
AC Q9HJ11; Q9NCW2;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Progressive ankylosis protein homolog (ANK).  
GN ANKH OR KIA1581.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20355194; PubMed=10894769;  
RA Ho A.M., Johnson M.D., Kingsley D.M.;  
RT "Role of the mouse ank gene in control of tissue calcification and  
RT arthritis." Science 289:265-270(2000).  
RL [2]  
RN [2]  
RP TISSUE=Brain;  
RX MEDLINE=20450683; PubMed=10997877;  
RA Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes.  
RT XVII. The complete sequences of 100 new cDNA clones from brain which  
RL code for large proteins in vitro." DNA Res. 7:273-281(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Ovary;  
RA MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Iqbal N.A., Peters G.J., Abramson R.D., Mullah S.J.,  
RA Bosa S.A., McKernan K.J., McKernan K.J., Walek J.A., Guntarine P.H.,  
RA Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butcherfield J.S.N., Krzywinski M.I., Skalska U., Skalius D.E.,  
RA Schermer A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RL human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP VARIANTS CMDU SER-375 DEL; PHE-376 DEL AND ALA-380 INS.  
RX MEDLINE=21313103; PubMed=11326338;

Sequence Alignment

RA Reichenberger E., Tiziani V., Watanabe S., Park L., Teki Y.,  
 RA Santana C., Baur S.T., Shiang R., Grange D.K., Belington P.,  
 RA Gardner J., Hamersma H., Sellars S., Ramesar R., Liddel A.C.,  
 RA Sommer A., Raposo do Amaral C.M., Gorlin R.J., Mulliken J.B.,  
 RA Olsen B.R.,  
 RA "Autosomal dominant craniofacial dysplasia is caused by mutations  
 in the transmembrane protein ANK.",  
 RA Am. J. Hum. Genet. 68:1321-1326(2001).  
 RN  
 RP VARIANTS CMDJ ARG-292; ARG-331; SER-375 DEL; PHE-377 DEL; ALA-380 INS  
 RP AND ARG-389.  
 RX MEDLINE=2125282; PubMed=11326272;  
 RA Nuerberg P., Thiele H., Chandler D., Hehne W., Cunningham M.L.,  
 RA Ritter H., Jeschik G., Uhlmann K., Mischung C., Harrop K.,  
 RA Goldblatt J., Borochowitz Z.U., Kotzot D., Westermann F., Mundlos S.,  
 RA Braun H.-S., Laining N., Tinschert S.,  
 RA "Heterozygous mutations in ANKH, the human ortholog of the mouse  
 progressive ankylosis gene, result in craniofacial dysplasia.",  
 RA Nat. Genet. 28:37-41(2001).  
 CC  
 CC -1- FUNCTION: Regulates intra- and extracellular levels of inorganic  
 CC pyrophosphate (PPI), probably functioning as PPI transporter.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- TISSUE SPECIFICITY: Found in osteoblasts from mandibular bone and  
 CC from iliac bone; not detected in osteoclastic cells.  
 CC -1- DISEASE: Defects in ANKH are the cause of craniofacial  
 CC dysplasia Jackson type (CMDJ) [MIM:123000]. CMDJ is a rare  
 CC autosomal dominant skeletal disorder characterized by abnormal  
 CC bone formation and mineralization in membranous as well as  
 CC endochondral bones. Progressive thickening of the bones can cause  
 CC narrowing of cranial foramina and can lead to severe visual and  
 CC neurological impairment, such as facial palsy and deafness.  
 CC -1- SIMILARITY: BELONGS TO THE ANKH FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF274753; AAF8039.1; -; -;  
 DR EMBL; AB046801; BAB13407.1; ALT\_INT.  
 DR EMBL; BC009835; AAH09835.1; -;  
 DR EMBL; BC014526; AAI14526.1; -;  
 DR Gene; HGNC:15492; ANKH.  
 DR MIM; 605145; -;  
 DR MIM; 123000; -;  
 DR GO; GO:0016021; C:Integral to membrane; IDA.  
 DR GO; GO:0019867; C:outer membrane; TAS.  
 DR GO; GO:0030504; F:inorganic diphosphate transporter activity; IDA.  
 DR GO; GO:0005315; F:inorganic phosphate transporter activity; IDA.  
 DR GO; GO:007626; P:locomotory behavior; NAS.  
 DR GO; GO:0030500; P:regulation of bone mineralization; TAS.  
 DR GO; GO:0001501; P:skeletal development; NAS.  
 DR Transport; Phosphate transport; Transmembrane; Disease mutation;  
 KM Deafness.  
 FT DOMAIN 1 85 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 86 106  
 FT DOMAIN 107 131 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 132 152  
 FT DOMAIN 153 158  
 FT TRANSMEM 159 179  
 FT DOMAIN 180 189  
 FT TRANSMEM 190 210  
 FT DOMAIN 211 326  
 FT TRANSMEM 327 347  
 FT DOMAIN 348 350  
 FT TRANSMEM 351 371  
 FT TRANSMEM 372 403  
 FT TRANSMEM 404 426  
 FT TRANSMEM 427 429  
 FT TRANSMEM 430 452  
 FT TRANSMEM 452 452

FT DOMAIN 453 492 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 292 292 W -> R (in CMDJ).  
 FT TRANSMEM 292 292 /FTID=VAR 012192.  
 FT TRANSMEM 331 331 C -> R (in CMDJ).  
 FT TRANSMEM 331 331 /FTID=VAR 012192.  
 FT TRANSMEM 375 375 Missing (in CMDJ).  
 FT TRANSMEM 375 375 /FTID=VAR 012194.  
 FT TRANSMEM 376 376 Missing (in CMDJ).  
 FT TRANSMEM 376 376 /FTID=VAR 012195.  
 FT TRANSMEM 377 377 Missing (in CMDJ).  
 FT TRANSMEM 380 380 /FTID=VAR 012196.  
 FT TRANSMEM 380 380 /FTID=VAR 012197.  
 FT TRANSMEM 389 389 G -> R (in CMDJ).  
 FT TRANSMEM 389 389 /FTID=VAR 012199.  
 FT TRANSMEM 78 78 N -> S (in REF. 1).  
 FT TRANSMEM 78 78 /FTID=VAR 012199.  
 FT TRANSMEM 492 AA; 54240 MW; 44BBEE08BDECB CRC64;  
 SQ SEQUENCE  
 Query Match 100.0%; Score 2527; DB 1; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-195; Indels 0; Gaps 0;  
 Matches 492; Conservative 0; Mismatches 0;  
 QY 1 MWKPEALTYWELIRFVPLVGLITNTAIDPGEQALNRGIAVVEDAVEMLASVGLAYSLMK 60  
 DB 1 MWKPEALTYWELIRFVPLVGLITNTAIDPGEQALNRGIAVVEDAVEMLASVGLAYSLMK 60  
 QY 61 FFGPMSDKNGVLYPVNKKRTRKAVLCMTVAGAAVFFLLAVSDVGYITNKLHHV 120  
 DB 61 FFGPMSDKNGVLYPVNKKRTRKAVLCMTVAGAAVFFLLAVSDVGYITNKLHHV 120  
 QY 121 DESVGSKTRARFLYLAAPFMDAMWTHAGILKKHYSFLVGCASISDVYIAOVFAILL 180  
 DB 121 DESVGSKTRARFLYLAAPFMDAMWTHAGILKKHYSFLVGCASISDVYIAOVFAILL 180  
 QY 181 HSHLEREPLIPILSLIWGALVRCCTTCLGYKYNHDIIPDRSGBELGADATIRKMSF 240  
 DB 181 HSHLEREPLIPILSLIWGALVRCCTTCLGYKYNHDIIPDRSGBELGADATIRKMSF 240  
 QY 241 MWPLLIATQRIISPIVNLFYSRDLGSSATTEVAITLTPYGVHMPYGLTERAVY 300  
 DB 241 MWPLLIATQRIISPIVNLFYSRDLGSSATTEVAITLTPYGVHMPYGLTERAVY 300  
 QY 301 PAFDKNPNKLVSTNTVTAAHKKFTFCVCAALTLCTCFWFTPTNYSKLLIDIGVD 360  
 DB 301 PAFDKNPNKLVSTNTVTAAHKKFTFCVCAALTLCTCFWFTPTNYSKLLIDIGVD 360  
 QY 361 PAFDKNPNKLVSTNTVTAAHKKFTFCVCAALTLCTCFWFTPTNYSKLLIDIGVD 360  
 DB 361 PAFDKNPNKLVSTNTVTAAHKKFTFCVCAALTLCTCFWFTPTNYSKLLIDIGVD 360  
 QY 421 LGVHGATGVGSLAGFVGESTMVAIAACVYRKKKKMENSATGEGDSMTMPTEE 480  
 DB 421 LGVHGATGVGSLAGFVGESTMVAIAACVYRKKKKMENSATGEGDSMTMPTEE 480  
 QY 481 VTDIVEMRENE 492  
 DB 481 VTDIVEMRENE 492  
 QY 481 VTDIVEMRENE 492  
 DB 481 VTDIVEMRENE 492  
 RESULT 2  
 ANKH MOUSE STANDARD; PRT; 492 AA.  
 AC Q9JH22; Q35138; Q35139;  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Progressive ankylosis protein (Fn54 protein).  
 GN ANKH OR ANK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]

Dp	2401	GGCTGCTTTTCCCTGCCTTCCTCCTTAAGGTGCATTAAGCGCAATCACATGAAGACATCC	2460
Oy	2461	TAACTTGCAATTATAGTTTTTATACAGTAACAATAAGCTTAAGTCCATACAGCATTTCTA	2522
Dp	2461	TAACTTGCAATTATAGTTTTTATACAGTAACAATAAGCTTAAGTCCATACAGCATTTCTA	2522
Oy	2521	TGCCAGGTTGCTGTAAAGGTAACATTGAAATAGATATATTAACCTGGTCTGGTATCCCTTA	2580
Dp	2521	TGCCAGGTTGCTGTAAAGGTAACATTGAAATAGATATATTAACCTGGTCTGGTATCCCTTA	2580
Oy	2581	GTCATTAACCTGGGTAACAGTAATGAGAATGTAACGTAACCTGCTCCACACCAT	2640
Dp	2581	GTCATTAACCTGGGTAACAGTAATGAGAATGTAACGTAACCTGCTCCACACCAT	2640
Oy	2641	ACGATTAAGCAACAACATTTTATPAAGATACAGACACTAGTGGTCTCTCTGAATA	2700
Dp	2641	ACGATTAAGCAACAACATTTTATPAAGATACAGACACTAGTGGTCTCTCTGAATA	2700
Oy	2701	ACGCATTGAAAATCCATGACATGACATATATTTCTAAGTTTGGAAAGCAGGTTTTT	2760
Dp	2701	ACGCATTGAAAATCCATGACATGACATATATTTCTAAGTTTGGAAAGCAGGTTTTT	2760
Oy	2761	CCCTTAAAAAATTTATAGCACAGGTTCACTAAATGATTAATGTCAGAAATCCTAGACTGA	2820
Dp	2761	CCCTTAAAAAATTTATAGCACAGGTTCACTAAATGATTAATGTCAGAAATCCTAGACTGA	2820
Oy	2821	AAGAACCTTAACAAAAAATATTTTAAAGATATAATATATGCTGATATATGTATGAT	2880
Dp	2821	AAGAACCTTAACAAAAAATATTTTAAAGATATAATATATGCTGATATATGTATGAT	2880
Oy	2881	TTATTTTAGGCTAATAACATTTCCATTTTGCATTTTGCATTAATAATGTCTTAATACA	2940
Dp	2881	TTATTTTAGGCTAATAACATTTCCATTTTGCATTTTGCATTAATAATGTCTTAATACA	2940
Oy	2941	AAAAA 2945	
Dp	2941	AAAAA 2945	
RESULT 3			
LOCUS	AB046801	3928 bp	mRNA linear PRI 22-FEB-2001
DEFINITION	Homo sapiens mRNA for KIAA1581 protein, partial cds.		
ACCESSION	AB046801		
VERSION	AB046801.1	GI:10047236	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	Nagase,T., Kikuno,R., Nakayama,M., Hirosewa,M. and Ohara,O. Prediction of the coding sequences of unidentified human genes. XVIII. The complete sequences of 100 new cDNA clones from brain DNA Res. 7 (4), 273-281 (2000)		
JOURNAL	MEDEV		
PUBMED	10997877		
REFERENCE	2 (bases 1 to 3928) Ohara,O., Nagase,T. and Kikuno,R. Direct Submission Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba 232-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913, Fax:81-438-52-3914)		
JOURNAL			
FEATURES	Location/Qualifiers		
Source	1..3928		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="fj05690"		
	/note="vector:pBluescriptII SK plus"		

gene 1. 3928  
/gene="X1AA1581"  
c1. 1640  
/gene="X1AA1581"  
/note="Start codon is not identified."  
/codon\_start=3  
/product="X1AA1581 protein"  
/protein\_id="BAB13407.1"  
/db\_xref="GI:10047237"  
/translation="APORGGGAAAAAPAPPAASVGPBPAPVAPRLPAPSPGRSR  
VWGQFAGTAMKFFPALTHYMPILRIYVPLSGTNTN1ADGBQALNGIAAYEDAYEM  
ASGYAIAIMKFEFSPMSDFKNVGLVFNPSRDKATLCHVYAGLAAVFTLLIAY  
IDGYIINKLHHVDSVGSKTRRAFLYLAAPFPMDAAVTAAGILLKRYSEFLVCA  
LSDVIAQYVAFTLLHSHLSKREPLLPIELSLVYGALRCYTLCLGYSKTHDILIP  
SGPELGGDAPTRKMLSEFWPLALILATORIRPRLVNFVPSDLGSSSAYTAVALIT  
TYPMHPVSGMLTEIRAYPAFPDKNBNBKNSTNSNTYTAHIKKEFTVMAALSLTLD  
FVMPNPVNSKERILIDIGVDYFPAELCVPLRISPEPVPVTVRAIRLTMGLMTLKK  
FLVAPSVATRIIVLASIVLVLYGAGHATLGVGSLIAGFVGBSTMAIACVYRKQ  
KKKNENESATGEDSANTMDMPTEHYVDIYMERENE"

Query\_Match 99.7%; Score 2936.6; DB 9; Length 3928;  
Beet Local Similarity 99.9%; Pred. No. 0;  
Matches 2939; Conservat. 0; Mismatches 4; Indels 0; Gaps 0;

Qy	1	CGCTGCGCCGCTGCGCCCTCGCCCTCCCGGAGAGTCCCTTGGCGAGCAGATGTGTG	60
Db	80	CGCTGCGCCGCTGCGCCCTCGCCCTCCCGGAGAGTCCCTCGCGAGCAGATGTGTG	139
Qy	61	GGGTGAGCCCACTGATGTGAATTCGCGCGCTGCAAGCACTCTGCGCCCT	120
Db	140	GGGTGAGCCCACTGATGTGAATTCGCGCGCTGCAAGCACTCTGCGCCCT	199
Qy	121	GATCGGCTTGGTGCCTCGCGGCAATCCATAGCATTCCTGGGAGAGAGGC	180
Db	200	GATCGGCTTGGTGCCTCGCGGCAATCCATAGCATTCCTGGGAGAGAGGC	259
Qy	181	CTTGAACCGGGGCACTCTGCTGTCAGAGAGATGCAAGTGAATGCTGGCCAGCTACG	240
Db	260	CTTGAACCGGGGCACTCTGCTGTCAGAGAGATGCAAGTGAATGCTGGCCAGCTACG	319
Qy	241	GCTGGCGTACTCCCTCATGAAGTCTTCAAGGGTCCCATGATGACTTCAAAAATGTGG	300
Db	320	GCTGGCGTACTCCCTCATGAAGTCTTCAAGGGTCCCATGATGACTTCAAAAATGTGG	379
Qy	301	CGGTGTTGTGAACGAGAGAGAGACAGACCAAGCCGTCCTGTATGTGTGTGC	360
Db	380	CGGTGTTGTGAACGAGAGAGAGACAGACCAAGCCGTCCTGTATGTGTGTGC	439
Qy	361	AGGGGCGCATGCGCGTCTTTCACACACTGATAGCTTATAGATTTAGATATACAT	420
Db	440	AGGGGCGCATGCGCGTCTTTCACACACTGATAGCTTATAGATTTAGATATACAT	499
Qy	421	TATCATTAATCTGACACATGTGAGCAAGTGTGGGAGACAGAGAGGCGCTTCT	480
Db	500	TATCATTAATCTGACACATGTGAGCAAGTGTGGGAGACAGAGAGGCGCTTCT	559
Qy	481	GTACTGCGCGCTTTCCTTTCATGAGACGGAATGGAATGAGCCCATGCTGGCATTTCT	540
Db	560	GTACTGCGCGCTTTCCTTTCATGAGACGGAATGGAATGAGCCCATGCTGGCATTTCT	619
Qy	541	AAAAACAATAACGTTTCTGGTGGATGTGCTCAATCTAGATGTATAGTCAGAGT	600
Db	620	AAAAACAATAACGTTTCTGGTGGATGTGCTCAATCTAGATGTATAGTCAGAGT	679
Qy	601	TGTTTTGAGCATTTTGTCTTCAAGTCACTGGAATCGCGGAGCCCTGTCTATCC	660
Db	680	TGTTTTGAGCATTTTGTCTTCAAGTCACTGGAATCGCGGAGCCCTGTCTATCC	739
Qy	661	GATCTCTCTGTGATAGGAGCACTTGTGGCTGACCACTGTGGCTGGGCTACCTA	720
Db	740	GATCTCTCTGTGATAGGAGCACTTGTGGCTGACCACTGTGGCTGGGCTACCTA	799

QY 721 CAAGAAATTCAGACATCATCCCTGACAGAAAGTGGCCCGAGAGCTGGGGGAGATGACAC 780  
 Db 800 CAAGAAATTCAGACATCATCCCTGACAGAAAGTGGCCCGAGAGCTGGGGGAGATGACAC 859  
 QY 781 AATAAGAAAGATGCTGAGCTTCTGGTGGCTTTGGCTTAAATTCCTGGCCACACAGAGAT 840  
 Db 860 AATAAGAAAGATGCTGAGCTTCTGGTGGCTTTGGCTTAAATTCCTGGCCACACAGAGAT 919  
 QY 841 CAGTGGCTATTTGTCAACCTCTTGTGTTCCCGGAGACCTTGGTGGCAGTTTGTGACGCAC 900  
 Db 920 CAGTGGCTATTTGTCAACCTCTTGTGTTCCCGGAGACCTTGGTGGCAGTTTGTGACGCAC 979  
 QY 901 AGAGGAGTGGCGAATTTTGAACGACACATACCTGTGGGTCAATGACATACGGCTGGT 960  
 Db 980 AGAGGAGTGGCGAATTTTGAACGACACATACCTGTGGGTCAATGACATACGGCTGGT 1039  
 QY 961 GAGGGAATCCGCTGTGTATCTCTGTTTGGACAGAAATPACCCGACAACTGTGT 1020  
 Db 1040 GAGGGAATCCGCTGTGTATCTCTGTTTGGACAGAAATPACCCGACAACTGTGT 1099  
 QY 1021 GAGCAGAGACACACAGTCAAGGAGCCCAATCAAGATTCACCTTGTGTGATGAC 1080  
 Db 1100 GAGCAGAGACACACAGTCAAGGAGCCCAATCAAGATTCACCTTGTGTGATGAC 1159  
 QY 1081 TCTGTCACTCAAGCTGT 1140  
 Db 1160 TCTGTCACTCAAGCTGT 1219  
 QY 1141 GATAGCATGATGGAGTGAATCTTGTGCTTGTGAACTGTGTGTGTGTGTGTGTGTGT 1200  
 Db 1220 GATAGCATGATGGAGTGAATCTTGTGCTTGTGAACTGTGTGTGTGTGTGTGTGTGT 1279  
 QY 1201 GTTCTCTTCTTCCAGATTCAGTCAAGTGAAGGAGCAATCAACCGGGGTGTGTGTGT 1260  
 Db 1280 GTTCTCTTCTTCCAGATTCAGTCAAGTGAAGGAGCAATCAACCGGGGTGTGTGTGT 1339  
 QY 1261 ACTGAAGAAACCTTGT 1320  
 Db 1340 ACTGAAGAAACCTTGT 1399  
 QY 1321 GAGCCTGT 1380  
 Db 1400 GAGCCTGT 1459  
 QY 1381 CCTGGCGGCTTGT 1440  
 Db 1460 CCTGGCGGCTTGT 1519  
 QY 1441 GAAAGCAAAAGAAAGATGAGATGAGTGGCCACGAGGGGAGAACTGTGCAATGAC 1500  
 Db 1520 GAAAGCAAAAGAAAGATGAGATGAGTGGCCACGAGGGGAGAACTGTGCAATGAC 1579  
 QY 1501 AGACATGTCTCGACAGAGAGATGACATGCTGTGAATGAGAGAGAGAAATGAATA 1560  
 Db 1580 AGACATGTCTCGACAGAGAGATGACATGCTGTGAATGAGAGAGAGAAATGAATA 1639  
 QY 1561 AGGCAAGGAGCGCATGAGGACCTGACAGGAGCGTCAAGTGAAGTACATCTGGCATAT 1620  
 Db 1640 AGGCAAGGAGCGCATGAGGACCTGACAGGAGCGTCAAGTGAAGTACATCTGGCATAT 1699  
 QY 1621 CTCTTCCCTCTCCCATGATATTTGTCCCTTTTGTGTTTGTGTTTGTGTTTGTGTTT 1680  
 Db 1700 CTCTTCCCTCTCCCATGATATTTGTCCCTTTTGTGTTTGTGTTTGTGTTTGTGTTT 1759  
 QY 1681 GGCCCTGATTAAGAGTTTGTGTGTATTTCTTACATATCTGGTATGCTCACTGACG 1740  
 Db 1760 GGCCCTGATTAAGAGTTTGTGTGTATTTCTTACATATCTGGTATGCTCACTGACG 1819  
 QY 1741 GGGGAGCTATGATGATGCTTTTACTGTGTATTAATAAACAAGAAACAATGACTT 1800  
 Db 1820 GGGGAGCTATGATGATGCTTTTACTGTGTATTAATAAACAAGAAACAATGACTT 1879  
 QY 1801 CATACCCCTGCTACGAAAAACCAAAAAGACAGCTGCTCAAGGTTGATGCTC 1860

Db 1880 CATACCCCTGCTACGAAAAACCAAAAGACAGCTGCTCAAGGTTGATGCTGCTC 1939  
 QY 1861 TCCCTCCCTGAGCAATCTCTCTTGAACCAAAAGACTGAGCTGTGACATGCGCTGCG 1920  
 Db 1940 TCCCTCCCTGAGCAATCTCTCTTGAACCAAAAGACTGAGCTGTGACATGCGCTGCG 1999  
 QY 1921 GTCACTCTGACAGAGGCGACAGACTCTCCGCTGCTCCCTCTTACGCTCTTAAAGATCA 1980  
 Db 2000 GTCACTCTGACAGAGGCGACAGACTCTCCGCTGCTCCCTCTTACGCTCTTAAAGATCA 2059  
 QY 1981 CAGGTTAAACTCGGCTTCTTGAATTTGCTTCCAGTCAATGAGCGTAAAGAGATG 2040  
 Db 2060 CAGGTTAAACTCGGCTTCTTGAATTTGCTTCCAGTCAATGAGCGTAAAGAGATG 2119  
 QY 2041 GAGCCCGGCTGCTCTTAAATTTCCCTTCCGACAGGAGTTGAAACATCTACCTC 2100  
 Db 2120 GAGCCCGGCTGCTCTTAAATTTCCCTTCCGACAGGAGTTGAAACATCTACCTC 2179  
 QY 2101 ACATGAGAGAGCGGCTGACGCTGAGCCCGGAGTCCCGCTTCACTGAGAAACGGA 2160  
 Db 2180 ACATGAGAGAGCGGCTGACGCTGAGCCCGGAGTCCCGCTTCACTGAGAAACGGA 2239  
 QY 2161 GACTGTGACACACAGAGCTGACAGATGACAGATCTCCGTAGAAAGTTGGTTG 2220  
 Db 2240 GACTGTGACACACAGAGCTGACAGATGACAGATCTCCGTAGAAAGTTGGTTG 2299  
 QY 2221 AAATGCGCGGGGAGAGAACTGACATGTTGAAATGATGATTCATCTGCGTCTC 2280  
 Db 2300 AAATGCGCGGGGAGAGAACTGACATGTTGAAATGATGATTCATCTGCGTCTC 2359  
 QY 2281 CTGATCTGAGAGAGCTGTAGTTCTCACCCCAACCGGTATATACATGAGCTAACTTT 2340  
 Db 2360 CTGATCTGAGAGAGCTGTAGTTCTCACCCCAACCGGTATATACATGAGCTAACTTT 2419  
 QY 2341 TTAATTTGTCAAAAGCGCATCTTCAGATTCAGACCCCTGACAGATCTTTCTGAA 2400  
 Db 2420 TTAATTTGTCAAAAGCGCATCTTCAGATTCAGACCCCTGACAGATCTTTCTGAA 2479  
 QY 2401 GGCCTGCTTCCCTGCGCTTCTGAAAGTGTGCTTGAAGAGTCAATGAGAGTCAATGAGATCC 2460  
 Db 2480 GGCCTGCTTCCCTGCGCTTCTGAAAGTGTGCTTGAAGAGTCAATGAGAGTCAATGAGATCC 2539  
 QY 2461 TAACTTTCATTTAGATTTTACAGTGAATGAAGCTTAAAGTCAATCCAGATTTCTAA 2520  
 Db 2540 TAACTTTCATTTAGATTTTACAGTGAATGAAGCTTAAAGTCAATCCAGATTTCTAA 2599  
 QY 2521 TGCAGAGTGTGTAGGATACCTTTGAAGTGAATATTAACCTGTTCTGCTATCTTA 2580  
 Db 2600 TGCAGAGTGTGTAGGATACCTTTGAAGTGAATATTAACCTGTTCTGCTATCTTA 2659  
 QY 2581 GTCAATACTCGGGGTACAGGTAATGAGATGATGATGATGATGATGATGATGATGATGAT 2640  
 Db 2660 GTCAATACTCGGGGTACAGGTAATGAGATGATGATGATGATGATGATGATGATGATGAT 2719  
 QY 2641 ACGATAAAGCAAGACATTTTAAACGATACAGAGTCAATGATGATGATGATGATGATGAT 2700  
 Db 2720 ACGATAAAGCAAGACATTTTAAACGATACAGAGTCAATGATGATGATGATGATGATGAT 2779  
 QY 2701 ACGATTCGAATTCATGACAGTGCATATTTTCTAAGTTTGTGAAACAGAGTTT 2760  
 Db 2780 ACGATTCGAATTCATGACAGTGCATATTTTCTAAGTTTGTGAAACAGAGTTT 2839  
 QY 2761 CCTTTAAAAAATTAAGAGCTTCACTTAATGATTTAGTCAAGATTCAGACTGA 2820  
 Db 2840 CCTTTAAAAAATTAAGAGCTTCACTTAATGATTTAGTCAAGATTCAGACTGA 2899  
 QY 2821 AAGAACCTAAACAAAAAATATTTTAAAGATATATATGCTGATATGATATGATAT 2880  
 Db 2900 AAGAACCTAAACAAAAAATATTTTAAAGATATATATGCTGATATGATATGATAT 2959  
 QY 2881 TTAATTTAGCTATTAATACATTTCTGATTTTGCATTTTCAATTAATGATGCTTAATCA 2940

Query Match	Best Local Similarity	98.4%; Score 2896.8; DB 9; Length 3013;
Matches 2912; Conservative	0; Mismatches 7; Indels 1; Gaps 1;	
QY 27	CCGAGAGTCCCTCGCGGAGAGAGTGTGTGGGTCAAGCCAGCGGGAGACTATAG 86	
DB 1	CCGAGAGTCCCTCGCGGAGAGAGTGTGTGGGTCAAGCCAGCGGGAGACTATAG 60	
QY 87	TGAAATTCGGGCGCTCAAGCACTACTGCCCCCTGATCCGATCTTGTGTGCGCCCTGAGCA 146	
DB 61	TGAAATTCGGGCGCTCAAGCACTACTGCCCCCTGATCCGATCTTGTGTGCGCCCTGAGCA 120	
QY 147	TCAACCAATAGGCATGACTTGTGGGGAGAGAGCGCTTGAACCGGGGCAATGCTCTCTGCA 206	
DB 121	TCAACCAATAGGCATGACTTGTGGGGAGAGAGCGCTTGAACCGGGGCAATGCTCTCTGCA 180	
QY 207	AGGAGATGCACTGCAAGTCTGGCCAGCTACAGGGCGTGAGTACCTCCATGAAGTCT 266	
DB 181	AGGAGATGCACTGCAAGTCTGGCCAGCTACAGGGCGTGACTCCCTCATGAAGTCT 240	
QY 267	TACGGGTCCCATGAGTGACTTCAAAATGTGGGCGCTGTGTGTGTAACAGCAAGAG 326	
DB 241	TACGGGTCCCATGAGTGACTTCAAAATGTGGGCGCTGTGTGTGTAACAGCAAGAG 300	
QY 327	ACAGGACCAAGGCGTCTGTATAGTGTGTGGCAGGGGCAATGCGTCCCTCTTACA 386	
DB 301	ACAGGACCAAGGCGTCTGTATAGTGTGTGGCAGGGGCAATGCGTCCCTCTTACA 360	
QY 387	CACGTATAGCTTATAGTATTTAGATTACTACATTAATCAATAAATGACACATGTGAGC 446	
DB 361	CACGTATAGCTTATAGTATTTAGATTACTACATTAATCAATAAATGACACATGTGAGC 420	
QY 447	AGTCGGTGGGAGCAAGACGAGAGGGCTTCTGTATACCTGCGCGCTTCTTTCAAG 506	
DB 421	AGTCGGTGGGAGCAAGACGAGAGGGCTTCTGTATACCTGCGCGCTTCTTTCAAG 480	
QY 507	ACGCATGATGATGAGACCCATGCTGGCACTTCTTTAAACCAAAATACAGTTTCTGTGG 566	
DB 481	ACGCATGATGATGAGACCCATGCTGGCACTTCTTTAAACCAAAATACAGTTTCTGTGG 540	
QY 567	GAGTGTCTCAATCTGATGTATATAGCTCAGGTTGTTTGTAGCACTTGTGCTTACA 626	
DB 541	GAGTGTCTCAATCTGATGTATATAGCTCAGGTTGTTTGTAGCACTTGTGCTTACA 600	
QY 627	GTCACCTGAATGCGGAGGCCCTGCTCATCCCAATCCTTCTGTATGATGGCGCAC 686	